electronic reprint

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449
Editors: **E. N. Baker and Z. Dauter**

Structures of two RNA octamers containing tandem G·A base pairs

Se Bok Jang, Katrien Baeyens, Mi Suk Jeong, John SantaLucia Jr, Doug Turner and Stephen R. Holbrook

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of two RNA octamers containing tandem G·A base pairs

Se Bok Jang,[a]* Katrien Baeyens,[b] Mi Suk Jeong,[a] John SantaLucia Jr,[c] Doug Turner[d] and Stephen R. Holbrook[e]

[a]Korea Nanobiotechnology Center, Pusan National University, Busan 609-735, South Korea, [b]Laboratory of Analytical Chemistry and Medicinal Physicochemistry, KU Leuven, E. Van Evenstraat 4, B-3000 Leuven, Belgium, [c]Department of Chemistry, 410 West Warren, Wayne State University, Detroit, MI 48202, USA, [d]Department of Chemistry, University of Rochester, Rochester, NY 14627, USA, and [e]Structural Biology Department, Physical Biosciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA

Correspondence e-mail: sbjang@pusan.ac.kr

Received 26 December 2003
Accepted 18 February 2004

PDB References:
rGGCGAGCC, 1sa9, r1sa9sf; rGICGAGCC, 1saq, r1saqsf

The crystal structures of two RNA octamers, 5′-GGC-(GA)GCC-3′ and 5′-GIC(GA)GCC-3′, have been determined from X-ray diffraction data to 2.8 and 2.7 Å resolution, respectively. The RNA octamers crystallize in isomorphous unit cells containing two mispairs arranged in a self-complementary manner and one single strand in the asymmetric unit. The single strand pairs with another single strand related by crystallographic symmetry to form a third unique double helix. Tandem non-Watson–Crick G·A/A·G base pairs of the sheared type comprise an internal loop in the middle of each duplex. The NMR structure of this octameric RNA sequence is also known, allowing comparison of the variation between the six crystallographic duplexes and the solution structure. In the symmetric duplex of the octamer containing inosine, the sheared G·A pairs incorporate a bound water molecule. This duplex also binds one water molecule per strand in the minor groove adjacent to the G·A pairs.

1. Introduction

The polymorphism of RNA base pairing introduces diversity into RNA structures and is utilized as the basis for specific RNA–protein, RNA–RNA and RNA–ligand recognition. Particularly for non-Watson–Crick base pairing, different hydrogen-bonding patterns of roughly equal stability may form depending on the surrounding base pairs, minimization of the overall free energy of the RNA and the solution conditions. Examples of this polymorphism are the pairs formed between guanine and adenine bases. Base pairing between guanine and adenine has been studied phylogenetically (Gautheret *et al.*, 1994), thermodynamically (SantaLucia *et al.*, 1990; Walter *et al.*, 1994; Serra *et al.*, 2002) and structurally by both high-resolution NMR (Wu & Turner, 1996; Wu *et al.*, 1997; SantaLucia & Turner, 1993; Burkard *et al.*, 1999; Heus *et al.*, 1997) and X-ray crystallography (Baeyens *et al.*, 1996). Base pairs between G and A have been structurally observed in model compounds (SantaLucia & Turner, 1993; Baeyens *et al.*, 1996), 5S ribosomal RNA (Correll *et al.*, 1997), the sarcin/ricin loop of 28S rRNA (Correll *et al.*, 1998), hammerhead catalytic RNA (Pley *et al.*, 1994; Scott *et al.*, 1995) and terminating GAAA tetraloops and their complexes with tetraloop receptors (Heus & Pardi, 1991; Cate *et al.*, 1996; Perbandt *et al.*, 1998; Leonard *et al.*, 1994). Four varieties of G·A base pairs have been observed previously: head-to-head or imino hydrogen-bonded G(*anti*)·A(*anti*) pairs both with (Correll *et al.*, 1997) and without a bridging water molecule (Leonard *et al.*, 1994), the sheared type of G(*anti*)·A(*anti*) base pairing in which the major-groove face of adenine

interacts with the minor-groove face of guanine (Baeyens *et al.*, 1996) and the unusual G(*syn*)·A$^+$(*anti*) base pairing (Pan *et al.*, 1999). In the case of tandem G·A pairs, the type of pairing has been shown to depend on both the order (G·A/A·G *versus* A·G/G·A; Wu *et al.*, 1997) and the surrounding base pairs (Wu & Turner, 1996; Heus *et al.*, 1997). The sheared type of G·A base pair preceded by a C·G pair binds a divalent cation with a geometry common to both an internal loop (Baeyens *et al.*, 1996) and hammerhead catalytic RNA (Pley *et al.*, 1994; Scott *et al.*, 1995).

Here, we present the crystal structures of two RNA octamers, rGGCGAGCC and rGICGAGCC, that form self-complementary duplexes in the crystal incorporating tandem G·A/A·G non-canonical base pairs in an internal loop.

2. Materials and methods

2.1. Synthesis, crystallization and data collection

The octaribonucleotide rGGCGAGCC was synthesized chemically (SantaLucia & Turner, 1993) and purified by anion-exchange (DEAE) FPLC column chromatography with a linear salt gradient from 0.4 to 2.0 *M* sodium acetate and a pH gradient from 50 m*M* Tris–HCl pH 6.8 to pH 7.3. Crystals of r(GICGAGCC) were grown at room temperature from a solution consisting of 1.5 m*M* RNA, 5 m*M* $MgCl_2$, 0.1 *M* HEPES pH 7.5, 4% PEG 400, 2 *M* $(NH_4)_2SO_4$ to dimensions of 400 × 200 µm, while r(GGCGAGCC) was crystallized from 1.3 m*M* RNA, 0.2 *M* $Zn(OAc)_2$, 0.1 *M* sodium cacodylate pH 6.5, 18% PEG 8K and crystals grew to 1000 × 70 µm. The crystals were mounted in thin-walled quartz capillaries and used for X-ray data collection on a Rigaku R-AXIS IIC imaging-plate system using Cu $K\alpha$ radiation (λ = 1.5418 Å) and φ scans with a scan width of 2.0° for 30 min. The data were processed with the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

2.2. Structure determination

The structures were determined by the molecular-replacement method using the *EPMR* program (Kissinger *et al.*, 1999). The solution structure of r(GGCGAGCC)$_2$ (SantaLucia & Turner, 1993) determined by NMR was used as a search model for the crystal structure of r(GGCGAGCC)$_2$ with pseudo-temperature factors calculated from the root-mean-square differences in the coordinates of the NMR models. Subsequently, the crystal structure of r(GGCGAGCC)$_2$ was used as the search model for r(GICGAGCC)$_2$. In both searches, the positioning of the first duplex was by a six-parameter search using data between 15.0 and 3.0 Å resolution. For r(GGCGAGCC)$_2$, the solution for the first molecule had a correlation coefficient of 0.502 (R = 55.1%) and the solutions of the second and third molecules had correlation coefficients of 0.580 (R = 51.4%) and 0.609 (R = 50.6%), respectively. For r(GICGAGCC)$_2$, the correlation coefficient was 0.405 (R = 55.8%) for the first molecule, which was used as a partial structure in the subsequent search for the other molecules. Adding the second duplex and the final single strand gave correlation coefficients of 0.472 (R = 52.3%) and 0.588 (R = 48.2%), respectively. Rigid-body refinement using *X-PLOR* (Brünger, 1992; Brünger *et al.*, 1997) reduced the R factor from 47.7% (R_{free} = 47.9%) to 47.3% (R_{free} = 47.3%) for r(GICGAGCC) and from 48.6% (R_{free} = 46.0%) to 48.0% (R_{free} = 45.6%) for r(GGCGAGCC) for data in the resoltuion range 10.0–3.0 Å. Both crystal structures had two and a half double strands (five single strands) in the asymmetric unit, with a third duplex being formed by a twofold symmetry axis acting on the lone single strand.

2.3. Model refinement

Data in the resolution range 8.0–2.8 Å were used for further refinement. Each cycle of refinement consisted of positional followed by simulated-annealing and finally B-factor refinement using the *CNS* program (Brünger *et al.*, 1998). Restraints were placed on bond lengths, bond angles, non-bonded contacts, temperature factors of neighboring atoms, planarity of the bases and non-crystallographic symmetry (NCS). Difference Fourier and $2F_o - F_c$ electron-density maps, as well as omit maps, were calculated at regular intervals to allow manual modification. The rebuilding of the model and the addition of solvent were performed using the *O* graphics program (Jones *et al.*, 1991). Solvent molecules were added conservatively with due regard for their environment, including potential interactions with hydrogen-bonding partners. The r(GGCGAGCC) and r(GICGAGCC) crystals belong to space group $P6_4$, with unit-cell parameters $a = b$ = 69.65, c = 68.36 Å, γ = 120° and $a = b$ = 69.65,

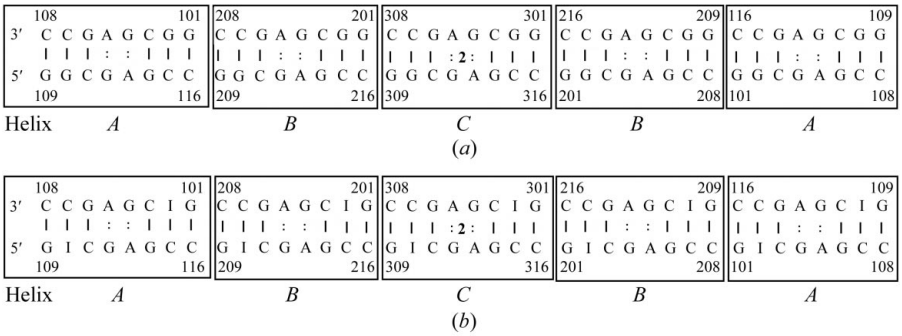


Figure 1
Schematic diagrams of (*a*) r(GGCGAGCC) and (*b*) r(GICGAGCC) as found in the crystal structures. The duplexes are boxed and the numbering of the 5′ and 3′ residues is indicated. The stacking pattern between the three crystallographically unique duplexes in each crystal is depicted by the adjoining boxes. The Watson–Crick base pairs are indicated by solid lines and the non-canonical G·A base pairs by double dots. The duplex formed by the octamer 301–308 is related to its pairing strand (309–316) by a crystallographic twofold axis represented by **2**. The stacking between the duplexes shows cross-strand stacking between the guanines. Pseudo-infinite helices pack in layers at 60° angles: one axis in the a* and one in the b* direction. At the junctions of the octamer duplexes there is cross-strand stacking between guanines. The figure shows the alignment of backbones. The layers interact by aligning the backbones and grooves.

c = 67.81 Å, γ = 120°, respectively. At the end of the refinement, the crystallographic R factor was 24.1% and the R_{free} value was 25.5% for r(GGCGAGCC)$_2$ with 20 bound water molecules. For r(GICGAGCC)$_2$ the final R factor was 23.8% and the R_{free} was 29.7% with 20 waters included. Electron densities were clear for all non-H atoms. The models both exhibit good geometry, with r.m.s. deviations from ideal bond lengths and angles of 0.006 Å and 1.074° for the r(GGCGAGCC) structure and 0.005 Å and 0.973° for the r(GICGAGCC) structure, respectively.

3. Results

In crystals of the RNA octamer rGGCGAGCC (GA8) and the isomorphous crystals of rGICGAGCC (IGA8) four unique self-complementary strands form two double helices, while a fifth strand forms a third duplex in the crystallographic asymmetric unit around a twofold crystallographic axis. Each of these six non-identical (three from GA8 and three from IGA8) duplexes incorporates an internal loop consisting of tandem G·A/A·G base pairs as shown schematically in Fig. 1. A representative duplex with its computed helical axis is shown in Fig. 2(*a*), together with a space-filling illustration in Fig. 2(*b*).

In all three GA8 duplexes and the two non-symmetric duplexes of IGA8, the G·A base pairs are of the sheared variety previously observed in the crystal structures of hammerhead catalytic RNA (Pley *et al.*, 1994; Scott *et al.*, 1995) and a synthetic RNA dodecamer (Baeyens *et al.*, 1996). These pairs are formed by three hydrogen-bonding interactions: N2(G)–N7(A), N3(G)–N6(A) and O2′(G)–N6(A). A fourth hydrogen bond has also been observed in solution between N2(G) and O1P of the adenine in the opposite strand (SantaLucia & Turner, 1993) for this sequence. In the three unique GA8 and IGA8 molecules observed in these crystal structures, we only observe cross-strand hydrogen bonding between N2(G) and the adenine phosphate for the symmetric G·A pairs of GA8 helix *C* (hydrogen-bonding distance 3.23 Å). A sheared G·A base pair forming three hydrogen bonds is shown in Fig. 3(*a*), while the G·A pair from GA8 helix *C* that makes four hydrogen bonds is shown in Fig. 3(*b*).

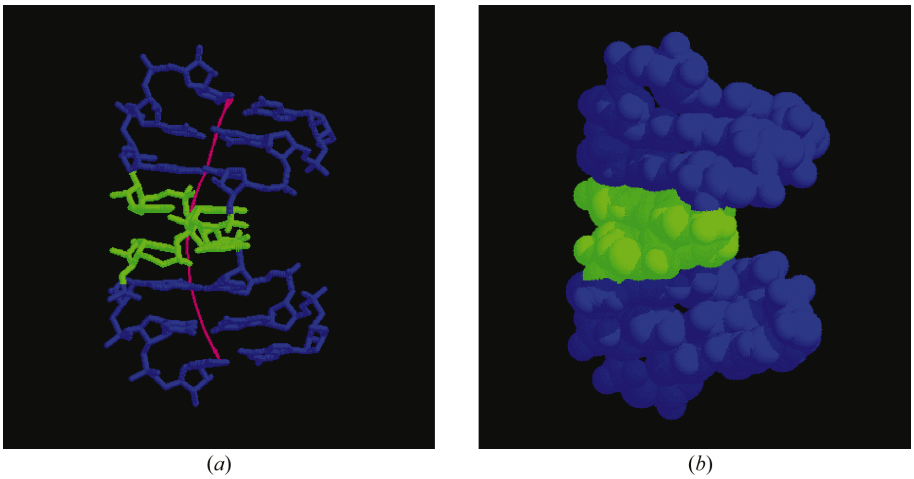
Table 1
Helical parameters of GA8 and IGA8: comparison with the average NMR structure (1yfv) and canonical A-RNA.

Groove widths and depths calculated at the G·A base pairs. Values in parentheses are standard deviations.

	Average helical twist (°)	Average rise per residue (Å)	Average major-groove width (Å)	Average minor-groove width (Å)	Helical diameter max/min (Å)	Helical curvature (Å)
GA8 *A*	32.73	3.62	7.23 (1.07)	9.45 (0.53)	18.4/16.4	39.7
GA8 *B*	31.54	3.55	8.62 (1.17)	10.81 (0.95)	19.1/15.9	41.8
GA8 *C*	32.65	3.55	7.22 (1.02)	9.26 (0.70)	18.4/15.6	38.4
IGA8 *A*	32.50	3.37	7.63 (1.30)	9.38 (0.55)	18.8/15.9	24.4
IGA8 *B*	30.57	3.53	9.02 (1.42)	11.23 (0.79)	19.2/16.4	35.6
IGA8 *C*	31.72	3.26	6.08 (0.01)	8.97 (1.32)	21.7/18.7	15.6
NMR	36.98	3.32	4.99 (0.06)	9.37 (0.94)	17.9/15.4	21.7
A-RNA	32.70	2.81	4.0	11.1	17.4/17.4	0.0

The six independent duplexes can be compared in several ways, including least-squares superposition, torsion angles and helical parameters. Global helical parameters for each of the six independent duplexes, the NMR solution structure (SantaLucia & Turner, 1993; PDB code 1yfv) and canonical A-form RNA, calculated with the *CURVES* program (Lavery & Sklenar, 1989), are compared in Table 1. These indicate a curvature of the helix axis by 23–35° induced by the tandem G·A mispairs, which are displaced by 2–3 Å from the axis. The local base-pair step parameters are shown in Table 2. The average helical twist is conserved by unwinding at the G·A/A·G step and overwinding between the base pairs adjacent to the mismatches. The results of pairwise least-square superposition of the duplexes are given as supplementary Tables 1 and 2.[1] From these tables, it is clear that the double helices of GA8 are as close in conformation to helices *A* and *B* of IGA8 as they are to each other and that all six helices are more similar to each other than to the NMR structure 1yfv. The most different of the crystallographic duplexes is helix *C* of IGA8. As discussed above, the two strands of helix *C* are related by exact crystallographic symmetry and incorporate bridging waters between the G·A base pairs.

The average pairwise r.m.s.d. for the GA8 structures is 0.83 Å; for the IGA8 structures it is 1.23 Å (0.8 Å excluding helix *C*) and it is 1.03 Å between GA8



(*a*) (*b*)

Figure 2
(*a*) Stick representation of the GA8 *A* duplex (Watson–Crick pairs in blue, G·A pairs in green) showing the curvature of the helical axis (pink) as calculated by the *CURVES* program. The curvature is into the major groove at the G·A base pairs. Bound waters are not shown in this illustration. (*b*) Space-filling representation of the GA8 *A* double helix shown in (*a*) using the same color coding and viewpoint.

[1] Supplementary data have been deposited in the IUCr electronic archive (Reference: HM5009). Details for accessing these data are given at the back of the journal.

Table 2
Least-squares comparison of GA8, IGA8 (crystal) and NMR (solution) duplex structures.

The averages/root-mean-square deviations are in bold (Å). The maximum deviations are shown immediately below in Å. For the NMR structure only the root-mean-square deviations are shown.

		GA8			IGA8		
		A	*B*	*C*	*A*	*B*	*C*
GA8	*A*	—	**0.89/1.00** 3.14 C16(O5′)	**0.49/0.60** 2.37 C16(O5′)	**0.44/0.53** 1.77 C15(O2′)	**0.71/0.81** 2.31 C16(O3′)	**1.43/1.51** 3.14 C16(O1P)
	B		—	**0.78/0.89** 2.50 A5(C2)	**0.91/1.00** 3.02 C16(C5′)	**0.99/1.10** 3.37 C16(C5′)	**1.25/1.33** 2.42 A13(C2)
	C			—	**0.57/0.63** 1.86 C8(O2′)	**0.79/0.84** 2.23 A13(C2)	**1.42/1.50** 2.86 A5(O2′)
IGA8	*A*				—	**0.71/0.80** 2.18 A13(N1)	**1.45/1.54** 2.7 C15(O2P)
	B					—	**1.29/1.36** 2.43 C15(O2P)
	C						—
NMR (1yfv)		1.16	1.54	1.29	1.17	1.47	2.15

and IGA8 (0.82 Å excluding helix *C*). While the average pairwise r.m.s.d. between NMR structures was only 0.7 Å, that between the X-ray crystallographic and NMR structure was 1.46 Å (1.33 Å excluding IGA8 helix *C*). This is a 62% greater difference between the crystal and NMR structures than between the crystal structures themselves (excluding IGA8 helix *C*).

The stacking between G·A base pairs is illustrated in Fig. 4(*a*) and the stacking of these non-canonical pairs with the neighboring Watson–Crick C·G pairs is shown in Fig. 4(*b*). Extensive cross-strand stacking between the adenine bases of the tandem G·A pairs is observed, with little interaction between the guanine bases.

Intermolecular interactions observed in the crystal packing are essentially of two types: (i) head-to-head stacking between the three octamer duplexes, resulting in pseudo-infinite helices throughout the crystal lattice, and (ii) hydrogen bonding interactions between helices, which occur almost exclusively through O2′ ribose hydroxyls (Tamura & Holbrook, 2002). The junctions between the pseudo-infinite helices are offset so that the guanine bases in the different helices stack and the guanine ribose lies below the cytosine base.

4. Discussion

The base pairs formed between guanine and adenine, G·A pairs, are of particular interest owing to their thermodynamic stability, widespread occurrence in biological molecules and polymorphism. The two ribo-octanucleotide crystal structures that we have determined are especially informative in providing six independent examples of tandem G·A pairs which can be compared with each other as well as with the structure of the same molecule determined by NMR methods.

Tables 1 and 2 and supplementary tables 1 and 2 compare the GA8 and IGA8 duplex structures amongst themselves and with the average NMR structure 1yfv. In Table 2, it is clear that helices *A*, *B* and *C* of GA8 and helices *A* and *B* of IGA8 are the most similar (r.m.s.d. = 0.53–1.10 Å, average r.m.s.d. = 0.82 Å). The five similar crystallographic duplexes are more similar to each other than to the NMR model 1yfv (r.m.s.d. = 1.16–1.54 Å, average r.m.s.d. = 1.33 Å). 1yfv was one of 11 model structures calculated by restrained molecular

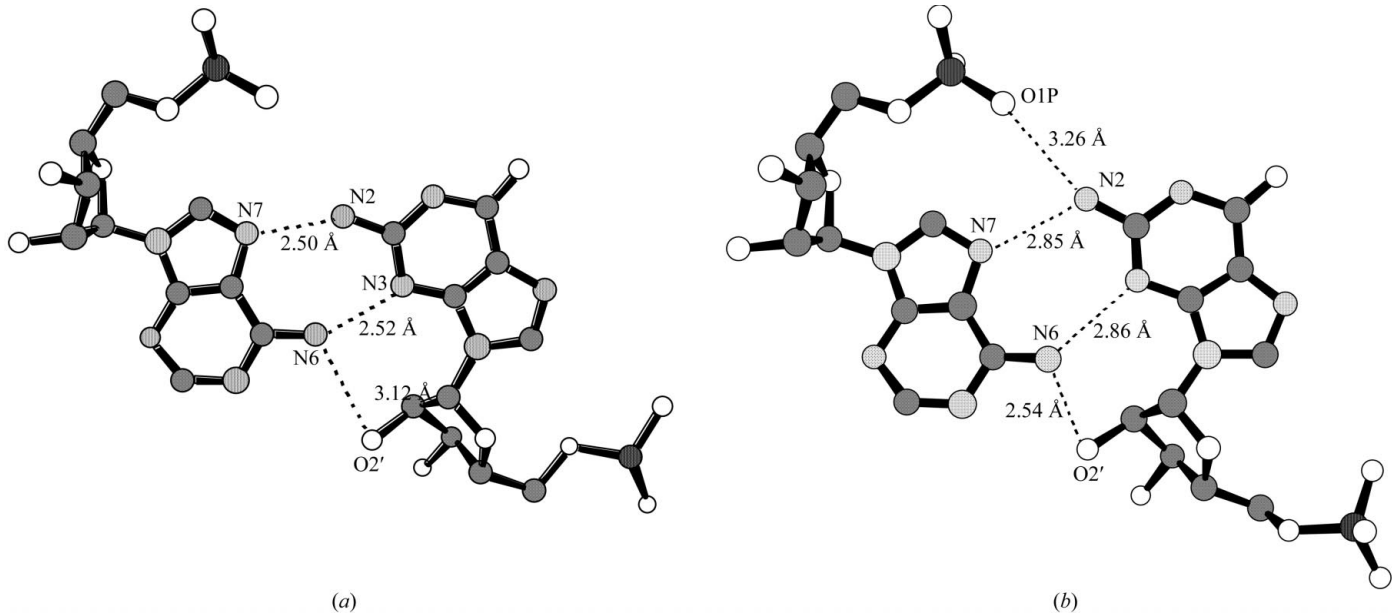
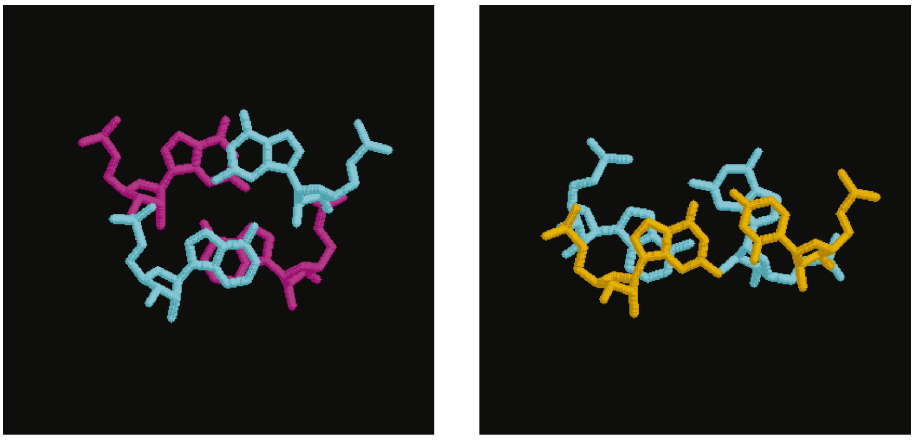


Figure 3
(*a*) The G·A sheared A(105)·G(112) base pair of GA8. N atoms are shaded lighter than C atoms and O atoms are unshaded. Three hydrogen bonds are indicated by dashed lines, with distances shown in Å. (*b*) The G·A sheared base pair A(305)·G(312) base pair of GA8. N atoms are shaded lighter than C atoms and O atoms are unshaded. Four hydrogen bonds are indicated by dashed lines, with distances shown in Å.



(*a*) (*b*)

Figure 4
(*a*) Stacking between the G·A (104, 113) pairs (cyan) and the A·G (105, 112) pairs (magenta) of GA8 helix *A*. (*b*) Stacking between the C·G (103, 114) pairs (gold) and the G·A (104, 113) pairs (cyan) of GA8 helix *A*.

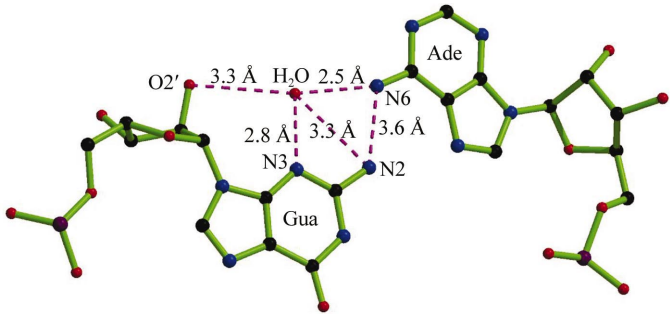


Figure 5
Structure of one of the symmetric G·A pairs in IGA8 helix *C* incorporating a bound water molecule. Distances in Å to the water are indicated.

dynamics that were consistent with the NMR data of 78 inter-proton distances, 18 hydrogen bonds (only for Watson–Crick pairs) and 26 dihedral angles per strand. The average r.m.s. deviation for the all-atom pairwise superposition of these 11 models is 0.70 Å, close to the internal consistency between the five similar duplexes of the crystal structures.

Helix *C* of IGA8 is the most different of the six independent duplexes found in the crystal structures. The r.m.s. deviation of helix *C* compared with the other five double helices ranges from 1.33 to 1.54 Å (average r.m.s.d. = 1.45 Å). The different conformation of helix *C* may be attributed to its bound water molecule and the bridging water between the G·A pairs found in this duplex (Fig. 5).

A more detailed comparison of the six octamer duplexes observed in the crystals and the solution (NMR) structure can be made by examination of their variable torsion angles as shown in supplementary Tables 1 and 2. The angles in bold show the largest deviations from canonical values. Generally, the Watson–Crick regions of the octamers show standard torsion angles, with the exception of the 3′-terminal cytosines. The major question is what changes in the dihedral angles are necessary to incorporate the sheared G·A base pairs into the double helix. Examination of supplementary Tables 1 and 2 shows that the ζ (O3′—P) torsion between the G·A and A·G base pairs differs from the standard *gauche*$^-$ angle ($\sim -60^\circ$) occurring in the Watson–Crick regions and 5′ and 3′ to the tandem G·A pairs. In the crystal structures, this conformational angle varies between -115 and -165°, with an average value of -139°. In the NMR structure 1yfv this angle is -108°, a smaller but still significant change from the canonical range. Interestingly, the ζ torsion angle is also -141.8° for the guanosine of the G·A base pair preceding an A·A pair in the crystal structure of the dodecamer duplex GGCCGAAAGGCC (Baeyens *et al.*, 1996). The other marked deviations in torsion angles are for α and γ of G104 and G304 in the GA8 crystal structure, and G212 of the IGA8 crystal structure. These angles change in a coupled manner from the standard g^-, g^+ torsions to *t*, *t* (*trans*, *trans*). This change is not observed in G204, G112 or G212 of GA8, the other IGA8 duplexes or 1yfv. Thus, there appears to be a secondary conformation that occurs in several of the duplexes but is not required for incorporation of the tandem sheared G·A pairs. Finally, there are a few anomalous ε (ribose C3′—O3′) torsion angles in the 180° range and one near 90° that are not conserved among the different duplexes.

A C·G/G·A motif in the hammerhead ribozyme has been observed to bind the divalent ions Mg^{2+}, Mn^{2+} and Cd^{2+} in the crystal (Pley *et al.*, 1994; Scott *et al.*, 1995; Wang *et al.*, 1999; Peracchi *et al.*, 1997) and magnesium bound at this site has been implicated to participate in the cleavage reaction. The same motif in an RNA dodecamer was shown to bind Mn^{2+} in a very similar manner (Baeyens *et al.*, 1996), coordinating with N7 of the guanine in the C·G pair and the phosphate of the adenine in the G·A pair. The octamers studied here have the same sequence motif, but no bound divalent cation is observed at the analogous position in these structures. The IGA8 structure was crystallized from a low magnesium concentration (5 m*M*) and a high ammonium sulfate concentration (2 *M*), while the GA8 octamer was crystallized from 0.2 *M* Zn^{2+} and low salt conditions. The presence of high salt coupled with low magnesium also resulted in no observable bound magnesium in the hammerhead ribozyme. The absence of observed zinc binding in GA8 may indicate inherent poor binding of that metal ligand by this motif.

An important question is whether divalent binding induces a conformational change in this sequence motif. Fig. 6 shows the superposition of a G·A pair from GA8 (helix *A*) and its C·G Watson–Crick neighbor with the C·G, G·A pairs from the RNA dodecamer duplex r(GGCCGAAAGGCC)$_2$ and its bound manganese ion. The GA8 *A* duplex is the most similar to the other duplexes of GA8 and IGA8 and is therefore a good representative. It is apparent from this superposition

that the C·G/G·A motif from the dodecamer is better positioned to bind manganese with the phosphorus ligand closer to the plane of the C·G pair and the guanine of the C·G pair translated toward the major groove. Both of the shifts are about 1 Å in magnitude. Although these differences can be attributed to the binding of a divalent cation, sequence effects cannot be ruled out since the following base pair is A·G in GA8 but is A·A in the dodecamer. The propagation of these changes to subsequent residues is also difficult to characterize without having the crystal structures of the identical sequence with and without bound cation.

Tandem mismatches in general and tandem G·A pairs in particular occur frequently in biological RNAs. An analysis of the occurrence of these non-canonical pairs in 16S and 23S rRNA indicates that G·A/A·G and A·A/A·G pairs are preferred and tandem A·G/G·A pairs are not observed. Structural analysis of A·G/G·A tandem pairs shows that they form head-to-head base pairs rather than the sheared pairs found here and in other G·A/A·G tandems. The ability of the sheared pairs to bind divalent cations and the resulting catalytic properties may be the critical factor in the biological preference for these tandems.

The G3·U70 wobble pair in the acceptor stem of *Escherichia coli* $tRNA^{Ala}$ has been shown to be necessary and sufficient for recognition by its cognate synthetase (Hou & Schimmel, 1988; McClain *et al.*, 1988). It has been shown that this G·U mispair can be substituted *in vivo* by other non-Watson–Crick base pairs, including G·A (and C·A), implying that distortion in the helix induced by mispairing rather than a particular sequence may be responsible for recognition (Gabriei *et al.*, 1996). The solution structure of the acceptor stem of this tRNA has been determined by NMR and demonstrates how a unique recognition site is provided by the wobble-base pairing (Ramos & Varani, 1997).

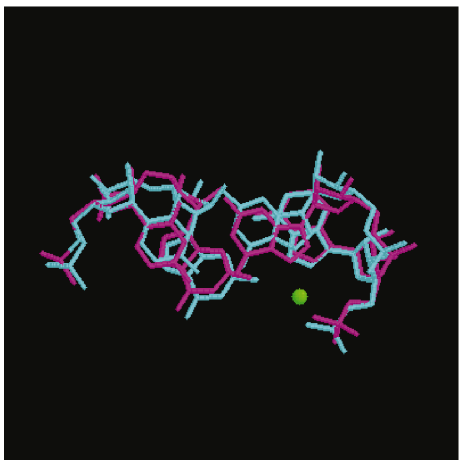


Figure 6
Superposition of the C·G/G·A base-pair motif from GA8 helix *A* (cyan) and an RNA dodecamer (magenta) containing a bound manganese ion in green. The guanine of the C·G pair in the dodecamer and the phosphate of the adjacent residue are relatively closer to the metal ion than the corresponding groups of GA8 helix *A*, in which no bound metal is observed.

Among the unique duplexes observed in these crystal structures, we have found sheared G·A pairs stabilized by three hydrogen bonds, four hydrogen bonds and by only a bridging water. This emphasizes the polymorphic nature of G·A pairing, as described in §1. This polymorphism suggests that G·A pairs could act as conformational switches which convert between different orientations of similar energy depending on environmental conditions (*i.e.* concentration of divalent cations).

We have determined the crystal structure of two octameric RNA sequences forming six unique duplexes in the crystal form that can be readily compared with the solution structure determined by NMR methods.

This work was supported NIH grant GM 4921501 to SRH and in part by Korea Research Foundation Grant KRF-2002-041-C00244 to SBJ. Facilities and equipment were provided through support of the Office of Energy Research, Office of Health and Environmental Research, Health Effects Research Division of the US Department of Energy.

References

Baeyens, K. J., De Bondt, H. L., Pardi, A. & Holbrook, S. R. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 12851–12855.
Brünger, A. T. (1992). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*. New Haven, CT, USA: Yale University Press.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Brünger, A. T., Adams, P. D. & Rice, L. M. (1997). *Structure*, **15**, 325–336.
Burkard, M. E., Kierzek, R. & Turner, D. H. (1999). *J. Mol. Biol.* **290**, 967–982.
Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. L. & Doudna, J. A. (1996). *Science*, **273**, 1696–1699.
Correll, C. C., Freeborn, B., Moore, P. B. & Steitz, T. A. (1997). *Cell*, **91**, 705–712.
Correll, C. C., Munishkin, A., Chan, Y.-L., Ren, Z., Wool, I. G. & Steitz, T. A. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 13436–13441.
Gabriei, K., Schneider, J. & McClain, W. H. (1996). *Science*, **271**, 195–197.
Gautheret, D., Konings, D. & Gutell, R. R. (1994). *J. Mol. Biol.* **242**, 1–8.
Heus, H. A. & Pardi, A. (1991). *Science*, **253**, 191–194.
Heus, H. A., Wijmenga, S. S., Hoppe, H. & Hilbers, C. W. (1997). *J. Mol. Biol.* **271**, 147–158.
Hou, Y.-M. & Schimmel, P. (1988). *Nature (London)*, **333**, 140–145.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
Lavery, R. & Sklenar, H. (1989). *J. Biomol. Struct. Dynam.* **6**, 655–667.
Leonard, G. A., McAuley-Hecht, K. E., Ebel, S., Lough, D. M., Brown, T. & Hunter, W. N. (1994). *Curr. Biol.* **2**, 483–494.
McClain, W. H., Chen, Y.-M., Foss, K. & Schneider, J. (1988). *Science*, **242**, 1681–1684.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pan, B., Mitra, S. N. & Sundaralingam, M. (1999). *Biochemistry*, **38**, 2826–2831.

Peracchi, A., Beigelman, L., Scott, E. C., Uhlenbeck, O. C. & Herschlag, D. (1997). *J. Biol. Chem.* **272**, 26822–26826.
Perbandt, M., Nolte, A., Lorenz, S., Bald, R., Betzel, C. & Erdmann, V. A. (1998). *FEBS Lett.* **429**, 211–215.
Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). *Nature (London)*, **372**, 68–74.
Ramos, A. & Varani, G. (1997). *Nucleic Acids Res.* **25**, 2083–2090.
SantaLucia, J. J., Kierzek, R. & Turner, D. H. (1990). *Biochemistry*, **29**, 8813–8819.
SantaLucia, J. J. & Turner, D. H. (1993). *Biochemistry*, **32**, 12612–12623.
Scott, W. G., Finch, J. T. & Klug, A. (1995). *Cell*, **81**, 991–1002.
Serra, M. J., Baird, J. D., Dale, T., Fey, B. L., Rotatagos, K. & Westhof, E. (2002). *RNA*, **8**, 307–323.
Tamura, M. & Holbrook, S. R. (2002). *J. Mol. Biol.* **320**, 455–474.
Walter, W. E., Wu, M. & Turner, D. H. (1994). *Biochemistry*, **33**, 11349–11354.
Wang, S., Karbstein, K., Peracchi, A., Beigelman, L. & Herschlag, D. (1999). *Biochemistry*, **38**, 14363–14378.
Wu, M., SantaLucia, J. J. & Turner, D. H. (1997). *Biochemistry*, **36**, 4449–4460.
Wu, M. & Turner, D. H. (1996). *Biochemistry*, **35**, 9677–9689.